MASTER'S THESIS

Evaluation Of The Effect Of Phased Guidance During Scripted Discussions Within Student Groups In Higher Vocational Education.

Kamp, Aaron

Award date:
2021

Evaluatie Van Het Effect Van Gefaseerde Ondersteuning in Script Gestuurde Discussies Binnen Studenten Groepen Binnen Het Hoger Beroepsonderwijs

Evaluation Of The Effect Of Phased Guidance During Scripted Discussions Within Student Groups In Higher Vocational Education

A.E. Kamp

Master Onderwijswetenschappen
Open Universiteit
Date: 12 April 2021
Name Supervisor: Dr. C. M. Stracke

Evaluation of the effect of phased guidance in scripted discussions on task-oriented reading within student groups in higher vocational education

Table of contents

Evaluation of the effect of phased guidance in scripted discussions on task-oriented reading within student groups in higher vocational education

Samenvatting

Evaluatie van het effect van gefaseerde ondersteuning in script gestuurde discussies op taak oriënterend lezen binnen studenten groepen binnen het hoger beroepsonderwijs

A.E. Kamp

Dit onderzoek zet het onderzoek van Darabi, Arrastia, Nelson, Cornille en Liang (2011) voort over gegroepeerde en gefaseerde discussies en het onderzoek van Wise, Saghafian en Padmanabhan (2012) over de mate van gefaseerde ondersteuning in script gestuurde discussie die optimaal is voor het aanleren van verbeterd leesgedrag.

Het doel van dit onderzoek is het bepalen van het effect van gefaseerde ondersteuning in script gestuurde discussies op het leesgedrag van studenten in het hoger beroepsonderwijs. Alle deelnemers aan dit onderzoek waren studenten aan de Hogeschool Rotterdam en werden geselecteerd op basis van hun deelname aan specifieke vakken die waren geselecteerd voor dit onderzoek. De studenten binnen de verschillende vakken werden willekeurig verdeeld in groepen en maakte gebruik van een technologie gestuurde leeromgeving (TELE) om opdrachten uit te voeren tijdens de onderwijsperiode van het vak. De helft van de groepen is aangeduid als de experimentele groep en voerde tijdens de training meerdere opdrachten in TELE uit gevolgd door discussies met gefaseerde ondersteuning, de andere helft van de groepen voerde de opdrachten in TELE uit doormiddel discussies zonder gefaseerde ondersteuning. Voor en na de training werd het leesgedrag gemeten door de studenten te monitoren tijdens een leestoets (leestijd van relevante en niet-relevante teksten voor elke vraag). Dit werd gemonitord met een speciaal voor het onderzoek van Okkinga (2018) ontworpen applicatie.

Er waren geen significante verschillen te vinden tussen het leesgedrag van de studenten met en zonder gefaseerde ondersteuning in script gestuurde discussies in dit onderzoek. Daarom kon tijdens dit experiment niet worden bepaald of gefaseerde ondersteuning in script gestuurde discussies een positief effect heeft op het leesgedrag van studenten. Er moet meer onderzoek worden gedaan om de optimale hoeveelheid gefaseerde ondersteuning in script gestuurde discussies te vinden voor het bevorderen van leesgedrag van studenten in het hoger beroepsonderwijs.

Sleutelwoorden: *Leesgedrag, Technologie ondersteunde leeromgeving, Gefaseerde begeleiding, Gescripte discussie.*

Summary

Evaluation of the effect of phased guidance in scripted discussions on task-oriented reading within student groups in higher vocational education

A.E. Kamp

This research builds on the research by Darabi, Arrastia, Nelson, Cornille, and Liang (2011) on grouped and phased discussions and on the research by Wise, Saghafian, and Padmanabhan (2012) on the amount of scripting that is optimal for learning reading behavior during a scripted discussion.

The aim of this research is to determine the effect of phased discussions on reading behavior of students in the higher vocational education. All participants in this study were students from the University of Applied Sciences of Rotterdam and were selected based on the subject. Students were randomly divided in groups and used a technology enhanced learning environment (TELE) to perform different assignments during the period of the subject. Half of the groups got the experimental training in which they used the TELE environment to have phased discussions about the reading assignments while the other half used the same TELE environment but with discussions that where not phased. Before and after the assignments the reading behavior was measured by monitoring their behavior during a reading test (reading time of relevant and non-relevant texts for the selected question). This was monitored using an application designed by Okkinga (2018).

No significant differences were found between the reading behavior of the students with and without phased discussions during this research. Therefore, in this experiment it could not be confirmed that phased discussions in TELE have a positive effect on the reading behavior of students. More research should be conducted to find the optimal amount of phased guidance during group discussions for students in the higher vocational education.

Keywords: *Reading behavior, Technology Enhanced Learning Environment, Phased guidance, Scripted discussion.*

Evaluation of the effect of phased guidance in scripted discussions on task-oriented reading within student groups in higher vocational education

1. Introduction

This research aimed to determining the amount of phased guidance during discussions that lead to improvement in reading behavior. To help students to improve on their reading behavior a training is given over several weeks that contains multiple reading assignments followed by a discussion in which one group is given instructions outright about the discussion and the other is helped in different phases during the discussion. In this section the problem sketch and the goal of the research are explained (section 1.1) followed by a brief review of the literature about reading behavior (section 1.2). The research question and hypotheses are introduced in section 1.3.

1.1. Problem sketch and goal of the research

Task-oriented reading is an important part of higher vocational education. Students need reading skills in order to succeed, but in higher vocational education those skills are rarely taught (Hermida, 2009). According to Hermida (2009), most students learn these reading skills in secondary education before they enter higher vocational education and they keep on using the same reading behavior. According to Mokhtari and Sheorey (2002), low-ability readers are not prepared to read academic reading materials when not guided with clear reading guidelines.

According to Vervoort and Elffers (2018), students from mbo (Dutch abbreviation for: *middelbaar beroepsonderwijs* = secondary vocational education) who attend hbo (Dutch abbreviation for: *hoger beroepsonderwijs* = higher vocational education) fail more frequently than students who come directly from the havo (Dutch abbreviation for: *hoger algemeen voortgezet onderwijs* = general secondary education) who attend hbo (21% vs. 12%). Students from mbo could be failing more often due to differences in reading behavior obtained in previous education as was previously stated but not conclusively proven by Elderman (2014). To increase the success for those students who attend higher vocational education (hbo), there is a need for extra teaching in the reading behavior and on reading in general as advised by the Expertgroep Doorlopende Leerlijnen Taal en Rekenen (2008).

Task-oriented reading is a method in which students know in advance that they need information from one or multiple texts: Task-oriented reading can help student to improve their reading behavior, independent of previous education, which they can use for the rest of their academic life (Vidal-Abarca et al., 2011). In this research we analyzed such a method to help students to improve their reading behavior in a class situation in hbo. We used a technology-enhanced learning environment to implement phased guidance in scripted discussions for the students since a teacher in the higher vocational education system cannot reasonably be asked to individually guide every student. The research aimed to analyze the effect of phased guidance in scripted discussions on the reading be-

havior of students. Two groups of students received a training with the same assignments: the experimental group was supported by phased guidance in scripted discussions while the control group had to do the scripted discussions without phased guidance.

The goal of this research project was to investigate the effect of phased guidance during scripted discussions about texts on subjects that the students were following for their specific curriculum. The technology-enhanced learning environment was designed to facilitate scripted discussions in two versions, with and without phased guidance. The reading behavior of the students was expected to be impacted as a result of whether or not they received phased guidance during the scripted discussions. When referred to phased or non-phased discussion in this thesis, it is meant that the information for the scripted discussion (roles, subjects, tasks) were given either in real time (= phased discussions) or completely at the start of the discussion (= non-phased discussions). The reading behavior was measured using a reading test before and after the training, which among others gave information about the time that the students took to read the text parts, to read the questions and to answer the questions. In this research we aimed to monitor the behavior of the students during the reading tests (pre- and post-training) to find out whether students who had been trained using the technology-enhanced learning environment showed different reading behavior based on the amount of phased guidance that they received during the scripted discussions.

1.2. Literature and theoretical background

In this section, I explain the different theoretical aspects of this research starting with a brief exploration on the literature on improving reading behavior followed by scripted cooperation.

1.2.1. Improving reading behavior

There are several relevant aspects to improve reading behavior during education. In this thesis, when referred to "reading behavior" the meaning as described by Harris and Hodges (1995) and elaborated on by Afflerbach, Pearson, and Paris (2008) is used. The reading behavior of a student means the way they read unknown texts to efficiently extract required information and comprehend and evaluate that information. Reading behavior is normally influenced by previous experiences with similar texts (Harris & Hodges, 1995). The relevant aspects to improve reading behavior is discussed below.

Instruction on reading behavior. According to Durkin (1978), there was little attention to the reading comprehension in elementary classrooms. But since the 1980's, elementary teachers increased the attention to reading comprehension of children (Pearson & Duke, 2002). Reading comprehension can be described as the ability to create a mental representation of a text (Kintsch, Patel, & Ericsson, 1999). Reading comprehension can be distinguished in lower and higher-level skills. The lower-level

skills include letter and word recognition, while higher level skills include the ability to give meaning to words and sentences (Perfetti, Landi, & Oakhill, 2005).

Reciprocal teaching. In reciprocal teaching, the teacher should guide students to a responsibility shift from teacher to student (Rosenshine & Meister, 1994). Reciprocal teaching consists of three instructional principles; (1) teaching comprehension-fostering reading strategies, (2) expert modelling and practicing, and (3) discussing reading strategies with other students (Brown, Palincsar, & Armbruster, 1984). All three principles should be guided by a teacher (Palinscar & Brown, 1984). Scaffolding (Gibbons, 2002) should be used to improve reading behavior within groups of students working together.

Effectiveness of reciprocal teaching. According to Spörer, Brunstein, and Kieschke (2009), reciprocal teaching appears to have a positive effect on the reading comprehension of students. Struggling readers profit even more from the effects of reciprocal teaching according to Edmonds et al. (2009). Rosenshine and Meister (1994) found that when applied to larger groups (more than 18 students) the positive effect is not significant in all studies used for their review study.

Whole-classroom context. Reciprocal teaching is designed for applications in small tutoring groups (Brown et al., 1984). It is found to be problematic to transfer reciprocal teaching to the classroom (De Corte, Verschaffel, & Van De Ven, 2001; Edmonds et al., 2009; Vaughn et al., 2013). Teachers found it hard to support students in gaining strategic reading skills and as a result, students showed poor application of reading strategies (Harlan, Bruce, & Edwards, 2016). While it is better to individually support students to improve their reading behavior, it is economically preferable for the higher vocational education system to teach in whole classroom context instead of teaching reading behavior to individual students (Westera, 2002). The use of reciprocal teaching asks for a different method of teaching because teachers will have to divide their attention to different student groups (Okkinga, van Steensel, van Gelderen, & Sleegers, 2018).

1.2.2. Scripted cooperation

In order to help teachers in guiding students in reciprocal teaching, cooperation scripts are advised (Kobbe et al., 2007). A cooperation script is a script which is expected to trigger specific interactions among students. According to Kobbe et al. (2007), a lot of scripts use a form of the Jigsaw method (Aronson & Bridgeman, 1979) where students form pairs with similar knowledge and the teacher provides them with complementary information to encourage interaction (Dillenbourg & Jermann, 2007).

One method of scripted cooperation is role distribution in which students get assigned different roles (Kollar, Fischer, & Hesse, 2006). Assigning specific roles has a positive effect on the levels of knowledge construction (Schellens, Van Keer, De Wever, & Valcke, 2007). The roles should be specific and get different instructions to trigger specific interactions and learning activities (Kobbe et

al., 2007). Kobbe et al. (2007) suggest the role of 'scientist' to promote planning, observing and drawing conclusions. By defining clear roles for each student, it is possible to prevent that students are behaving in a way that is linked to a role they prefer during assignments: this will likely happen when roles are not defined and can lead to certain students acting in a similar way or role for each assignment (Dillenbourg, 2002; Soller, Goodman, Linton, & Gaimari, 1998). When not properly distributed and defined, roles will be enacted depending on the situation and preference of the group, as a result they are unlikely to change during the course (Strijbos & Fischer, 2007). According to Strijbos and Weinberger (2010), defining a 'group leader' which is assigned to lead the group discussion and a 'writer' to summarize the findings will help to improve the cooperation.

The use of a computer ensures that without direct guidance from a teacher, collaboration can be encouraged within groups of students (Vogel, Wecker, Kollar, & Fischer, 2017). By using a whole system, it is possible to motivate the students to read, but also to help each other interpreting the texts they read (Dillenbourg & Tchounikine, 2007). A computer-supported collaboration script can be run as a part of a bigger system to help students in reading and interpreting texts without the guidance of a teacher (Dillenbourg & Tchounikine, 2007).

1.3. Research questions and hypotheses

The central question in this research is: "Do phased guidance in scripted discussions improve the individual reading behavior of students in higher vocational education?".

To answer the research question, the reading behavior of students is monitored and measured in two reading tests (before and after the training). The reading tests consist of questions and texts that contain information needed to answer the questions. Between the two tests, the students attended a training in which they had to complete given group assignments as described in section 2.5, either with phased guidance in scripted discussions about the assignments or without phased guidance in scripted discussions.

The assumptions are that students who have had phased guidance in the scripted discussions during the training improved their reading behavior more than students without phased guidance in scripted discussions during the training. Consequently, the students with phased guidance in scripted discussions during the training open less irrelevant texts before answering a question, study relevant text longer and correctly answer questions more frequently than students without phased guidance in scripted discussions during the training.

The five research questions are as follows;

1. RQ1: Are students who have had phased guidance in scripted discussions during the training, opening less texts in general before answering a question in the post-training reading test than students without phased guidance in scripted discussions during the training?

2. RQ2: Are students who have had phased guidance in scripted discussions during the training, using more time to read texts before answering a question in the post-training reading test than students without phased guidance in scripted discussions during the training?
3. RQ3: Are students who have had phased guidance in scripted discussions during the training, using more time to read relevant texts before answering a question in the post-training reading test than students without phased guidance in scripted discussions during the training?
4. RQ4: Are students who have had phased guidance in scripted discussions during the training, reading less irrelevant texts before answering a question in the post-training reading test than students without phased guidance in scripted discussions during the training?
5. RQ5: Are students who have had phased guidance in scripted discussions during the training, answering the questions on the post-training reading test better than students without phased guidance in scripted discussions during the training?

The five hypotheses for the research questions are as follows;

1. To compare the average number of relevant texts opened, there should be no significant difference between the total number of texts opened on the pre-training reading test by either the control- or the experimental group. If the training effected the reading behavior more positively for the students who have had phased guidance in scripted discussions during training, it can be expected that the students need less texts before answering a question on the post-training reading texts. Therefore, the hypothesis is as follows:
   H1: Students who have had phased guidance in scripted discussions during the training will open less texts before answering a question on the post-training reading test than students without phased guidance in scripted discussions during the training.
2. To compare the average time used to read relevant texts, there should be no significant difference between the total time used to read texts on the pre-training reading test by either the control- or the experimental group. If the training effected the reading behavior more positively for the students who have had phased guidance in scripted discussions during the training, it can be expected that the students need less time reading texts before answering a question on the post-training reading texts. Therefore, the hypothesis is as follows:
   H2: Students who have had phased guidance in scripted discussions during the training will use less time reading texts before answering a question on the post-training reading test than students without phased guidance in scripted discussions during the training.

3. If the training effected the reading behavior more positively for the students who have had phased guidance in scripted discussions during the training, it can be expected that the students need less time before finding relevant texts in the post-training reading test, since they will use less time before they realize that a certain texts is not relevant for the specific question. Therefore, the hypothesis is as follows:
   H3: Students who have had phased guidance in scripted discussions during the training will spent more time reading relevant texts during the post training reading test than students without phased guidance in scripted discussions during the training.
4. If the training effected the reading behavior more positively for the students who have had phased guidance in scripted discussions during the training, it can be expected that the students can find the location of the information in a text on the post-training reading test without opening a lot of other texts due to their reading behavior improved during the training. Therefore, the hypothesis is as follows:
   H4: Students who have had phased guidance in scripted discussions during the training will read less irrelevant texts before answering a question on the post-training reading test than students without phased guidance in scripted discussions during the training.
5. If the training effected the reading behavior more positively for the students who have had phased guidance in scripted discussions during the training, it can be expected that the students can find the information needed to answer the question sooner and can process this information better. Therefore, the hypothesis is as follows:
   H5: Students who have had phased guidance in scripted discussions during the training answer questions on the post-training reading test better than students without phased guidance in scripted discussions during the training.

2. Method

2.1. Design

The researchers searched for courses in which students had to read a lot and which were given to students during their curriculum within the Rotterdam University of Applied Sciences. At the Rotterdam University of Applied Sciences, a course consists of multiple gatherings for the duration of 8 weeks in which a specific subject is thought by a teacher and is part of the 4-year curriculum of the student. The teachers of the courses that were selected for the experiment have been asked which texts students had to read during the course. If students have to read a lot of texts which are relatively difficult to interpret, their teachers were asked if they would like to join the research program in order to use the specific training and the technology enhanced learning environment (TELE) to guide the students during the reading assignments during the course. It was not allowed for the participating teachers to use findings from the training or reading tests for examination purposes.

The students which choose to join the research-based course were given a tablet for the duration of the course in which they could work on the assignments during the training. On the tablet they could use an application which was designed especially for this research project and was adapted to fit each course with texts and assignments delivered by the participating teachers. The application is described in the section 2.4 Materials below.

The research project consisted of the training and two reading tests before and after the training which measured the time that students have spent on reading a question and searching for specific text parts during the reading test to measure the reading behavior of individual students before and after the training using the TELE. This measurement was done in the first and in the last session of the course, between the sessions all the students use the TELE with continuously changing roles. The students were assigned to different discussion groups of 3 to 5 students and each discussion group was placed in either the experimental group with phased group discussions in the TELE or the control group without phased group discussions but with a more general instruction during the discussion within the TELE. The teacher was equally available for all discussion groups and was not informed on whether the discussion groups where in the control- or the experimental group. Based on the described phases of groupwork, different phases had been described for the students, as presented in Appendix 3. All the discussion groups received the same amount of information but for the discussion groups within the experimental group the information was given in phases during the group discussions.

The time spent for reading questions and text parts during both the pre- and post-training reading test is compared between the experimental and control group and will be discussed in the section

3. Results. In this master thesis, only the averages for each student (average over all answered questions for each student in either the pre- or post-training reading test) for the following indicators will be used for answering research questions (with abbreviations for readability in the results section):
($\overline{n_t}$) = *Average number of texts read* – Average number of occurrences over all questions that a text is read prior to answering a question. Whenever a question is opened and closed every text that is opened before opening a new question is counted as a read text for that specific question for that student. The average over the stored number of texts for all questions for that student on that test is taken. This is used for answering RQ1.
($\overline{t_t}$) = *Average total reading time on the reading test* – The average time span over all questions that a text was read for a specific question during answering the questions on the reading test. Whenever a question is opened and closed, the time the text is opened is measured and stored for that specific question for that student. The average over the stored times for all questions for that student on that test is taken. This variable is used to answer RQ2.
($\overline{t_r}$) = *Average time reading relevant texts on the reading test* – The average time span over all questions that a relevant text for a specific question is opened during answering the questions on the reading test. Whenever a question is opened and closed, every text that is opened before opening a new question is either relevant or irrelevant for the previously read question, if relevant the time the text is opened is measured and stored for that specific question for that student. The average over the stored times for all questions for that student on that test is taken. This variable is used to answer RQ3.
($\overline{n_{nr}}$) = *Average number of non-relevant texts opened during the reading test* – The average number of occurrences over all questions that a non-relevant text for a specific question is opened during answering the questions on the reading test. Whenever a question is opened and closed every text that is opened before opening a new question is either relevant or irrelevant for the previously read question, if irrelevant the number of irrelevant texts opened for that question is increased by one for that specific question for that student. The average over the stored number of irrelevant texts for all questions for that student on that test is taken. This variable is used to answer RQ4.
(r_m) = *Reading test score* – Fraction of all questions that are answered correctly by a specific student on the reading test. This is calculated by dividing the test score by the maximum score possible for the reading test. This is used to answer RQ5.

2.2. Participants

In this research there were 104 participants over a period of 15 months. The participants were students from the Rotterdam University of Applied Sciences who were following a course that leads to different mayors (e.g., Biology, Physics, etc.). The students were spread over five different subjects: Physics

and ICT, pharmacology, educational sociology, research practice and government finances. The students of the subject's physics and educational sociology were in the second year of their total four-year educational course, while the students of the other three subjects were in their third year of their total four-year educational curriculum.

There were five different teachers from each subject to help the pre-service teachers during their assignments. Next to the author, there were two more researchers involved in the research (dr. Okkinga and dr. van Gelderen). Okkinga was the lead researcher during the experiment and was responsible for the data collection and reviewing the answers of the students on the reading tests. Van Gelderen was the supervisor that had an advisory role for this research project. The students were grouped by their own choice and the groups were randomly placed in either the control or the experimental group. The author of this thesis joined the research group after the first experiment and data collection. He participated during the second experiment and data collecting (during his physics classes) that is the basis of this master thesis. The author supported the processing of the information by writing and implementing the programming scripts for automatic data analysis. Therefore, the overall method could not be changed during this project to keep comparability between the different majors.

2.3. Technology-enhanced learning environment

In this section I will explain different choices that have been made and I will introduce literature on which these choices have been based. I start with an explanation on the type of assignments followed by the order of these assignments. I will end with a brief explanation of the different phases of group work that can be found in literature.

Types of assignments. To encourage learning and group discussions, assignments are needed that promote reciprocal learning (Stegmann, Weinberger, & Fischer, 2007). The assignments will contain parts that encourage discussion and therefore, do not always have a directly identifiable right or wrong answer. The components that were chosen for this research are: answering study questions (Hu & Li, 2017), summary assignment (Kirkland & Saunders, 1991), making mind map (Wolf, 2003), comparing texts (Kim & Lombardino, 2015), coming up with exam questions and an application assignment (Kim & Lombardino, 2015).

Order of assignments. In order to properly build up these group discussions, it is necessary to build up tasks that are increasing in complexity as described by Commons, Trudeau, Stein, Richards, and Krause (1998), only in this way the desired scaffolding can take place according to Pea (2004). The order of assignments also appears to be more important for more complex class situations, such as a group discussion (Davis & Miyake, 2004). Building on previous research done by Campbell (1988), Liu and Li (2012) have created a framework with which assignments can be classified according to complexity. In this training, particular attention was paid to the following elements of complexity:

number of information indications (Zhang, Li, Wu, & Wu, 2009), number of possible outcomes (Darisipudi, 2006), diversity of the elements (Ham, Park, & Jung, 2011), cognitive requirements (Bailey & Scerbo, 2007), required prior knowledge (Bailey & Scerbo, 2007), recognizability of task (Harvey & Koubek, 2000). Based on the described conditions, a classification has been made by Okkinga (2018). It is likely that not all assignments will be of an ascending order of complexity for all students, but because all the students make the same assignments it is likely that this will affect all students equally.

Phases of groupwork. When working in a group, it is advisable to offer good guidance during the discussion (Oliveira, Tinoca, & Pereira, 2011). When guidance is absent, students can get stuck or do not properly perform their role (Alavi & McCormick, 2008). Therefore, to properly guide the discussions there is a need for a more extensive and phased instruction in which the role for the student is explained in real time during the discussion (Darabi et al., 2011). In the description and guidance of the roles, there is a danger of giving too much guidance which will have a negative effect on the discussion outcome (Wise et al., 2012). The proper amount of guidance during the discussion is not yet clear, no conclusive research could be found. The goal of this research is to determine if an effect can be found in reading characteristics of students based on whether the students received phased scripted guidance during the discussions in training or they did not receive the phased scripted guidance. If a difference can be found, further research could be conducted to find the optimum amount of phased guidance during scripted discussions for the training on reading behavior.

2.4. Materials

The students were given a reading test with 11 questions for both, the pre- as well as the post-training reading test. During the two reading tests, different aspects of reading behavior were measured (e.g., amount of relevant texts read, time used to read, etc.) as the application is designed to measure the time spent on a certain screen during the reading tests. To make sure that the student was actually looking at the said part of the text, the other parts were blurred for the student, see figure 1 for an example of a screen during the reading tests.

Figure 1: Example of blurred texts during a reading test.

The reading tests generated a file for each student in which every action performed by the student during the reading test was logged; the time when students opened a part of the text and when they closed it, the time when students opened a question and when they closed the question. Parallel to the output of the application, there is a file which is created by the researcher (Okkinga) containing information about a question and which of the texts were relevant for that specific question, to be used in order to determine if a text read by a student was relevant for the answering of the last opened question or not (both files used in this research are presented in appendix 1). Students could freely move between different questions, even after an answer was saved, so the answers could be modified. The goal was to monitor the different indicators for each student and each question in the pre- and post-training reading tests. Not all indicators are used for the scope of this master research, the variables used are explained in section 2.1. Design. It is important to note that for each student the reading behavior is measured for each separate question during both reading tests, but in this master thesis only the averages over all questions are used.

The data from the pre- and post-training reading tests are compared to find out whether there is a difference between the control and the experimental groups. To extract the data from the application, a script was created by the author to extract information from the output of the application. There is no difference between the reading tests for students in the control or experimental group, but both tests (pre- and post-training reading test) contained different questions and texts and therefore, the two tests cannot be compared directly. To compare if there is a change in the reading behavior as a result

of the phased scripted discussions during training between the two reading test, the differences between the experimental- and the control-group are compared using a t-test. The t-test is performed on the pre-training reading test as well, this is done to check if the reading behavior does not differ between the students in the control- and experimental group.

The difference in training between students in the control or experimental group was the phased scripted discussion within the application during the reading assignments in the training: The discussion was either phased or not phased (Chesley, Cheng, Seban, & Landau, 2005). As said, the reading tests themselves (before and after the training with different reading assignments) were the same for both groups (control and experimental group) and were done individually. The training was done within the discussion groups and these groups remained the same for the entire course. Both, the experimental and control group used the same Technology Enhanced Learning Environment and all formed discussion groups (3 to 5 students) followed regular classes and received the same reading assignments. The difference was that the control group directly received the complete assignment with roles and descriptions, as can be seen in figure 2, while the experimental group was guided in phases during the discussion as presented in figure 3, 4 and 5. All groups received the same information at the end of every task. Therefore, the total information given during the assignments and group discussions remained the same, but during the discussions the experimental group received the information about the discussion in phases. The exact data given to students can be found in appendix 3.

Figure 2:Example of a screen shown to students in the control group during a reading assignment.

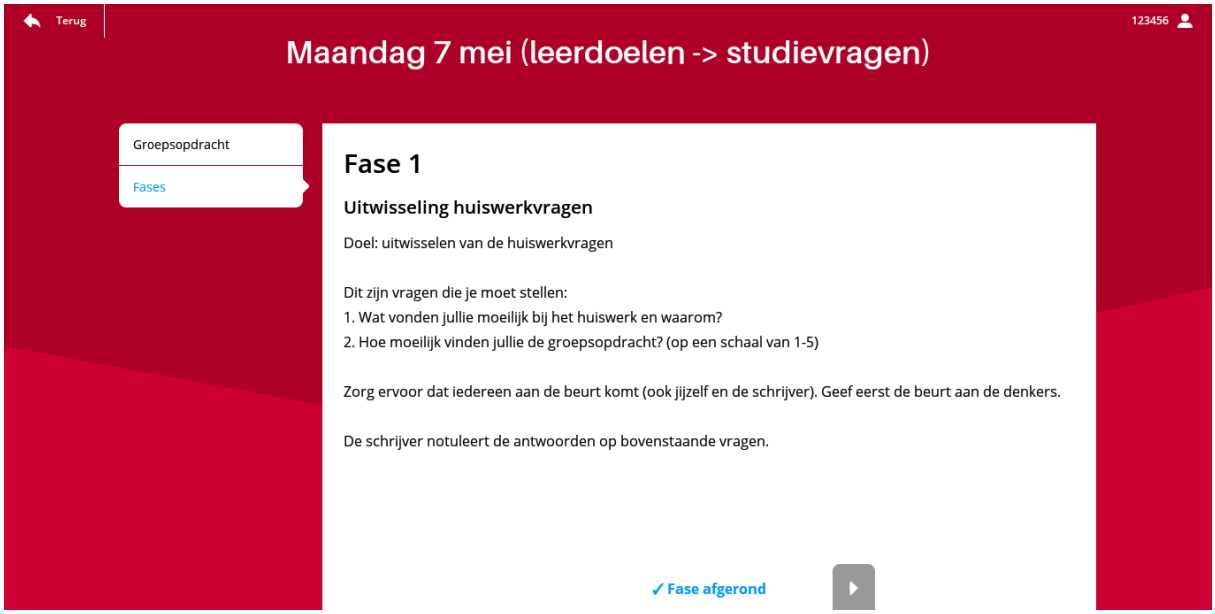

Figure 3: First of the five phases in which the experimental groups discuss the text.

Every assignment during training is done within a certain role in each discussion group and there were three roles based on Strijbos and Weinberger (2010): chair, writer or thinker. The students in the discussion group are assigned to one role and see different screens and receive different information within the application. The chair is the leader of the discussion and receives information as a help to lead the discussion. The writer receives the task to write down the conclusions made by the discussion group during the discussions: this is done by answering specific questions about texts the students had to read for the assignment in the TELE. The questions are answered within the application and the answers are automatically sent to the teacher. They are anonymously provided so that no names could be seen and used by the teacher. The answers should guide the teacher during the next lesson that is led by the teacher and is not part of the training. The thinker is the only role that can be performed by multiple students in a single group during the group discussions: Thinkers receive the task to answer the questions asked by the chair. They also receive their own information screen to help with their role, as presented in figure 4. The screen that is shown to a student depends on the role of the student and whether the discussion group is either an experimental or control group. If there are two thinkers within a single discussion group, they will receive the same screen. The other roles (leader and writer) are encouraged to think and contribute, but they are requested, to prioritize their own role over the contribution to the group discussion.

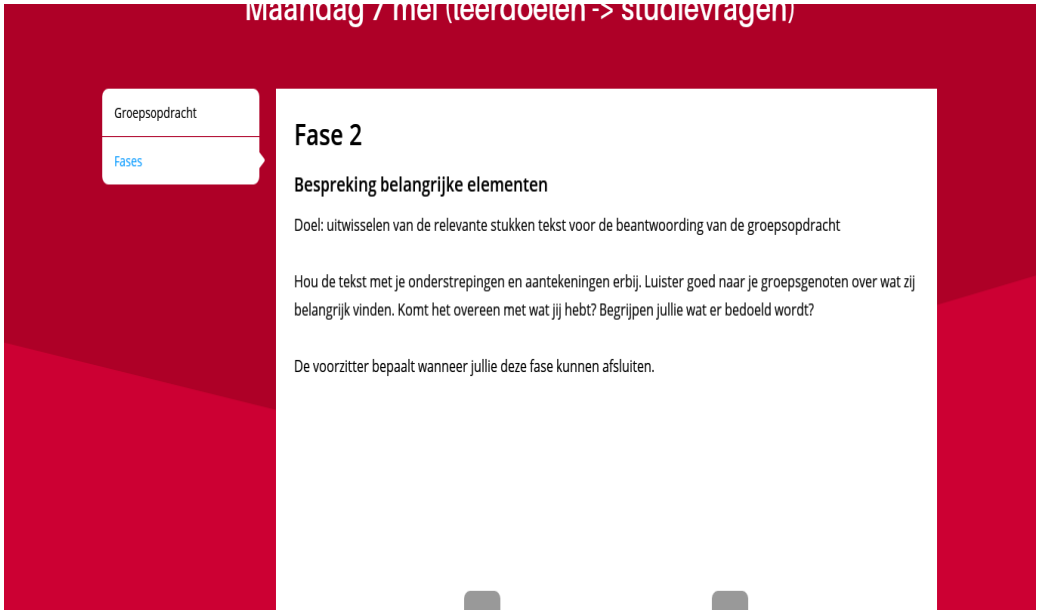

Figure 4: Second of the five phases in which the experimental groups discuss the text.

All discussion group members of both the control- and the experimental group receive the same information, but the members of discussion groups which are in the control group receive the information trough the TELE in one single moment whereas the members of discussion groups which are in the experimental group receive the information through the TELE in the following five different phases to improve the scripted discussions with real-time phases (based on Okkinga, 2018):

Phase 1: Orientation on the task - In the first phase, the discussion group is supported through the orientation on the task, the main approach to help the discussion group members in the first phase is by asking them questions about the homework reading assignment they had to do.

Phase 2: Discuss all relevant information for the group task – In the second phase, the discussion group is asked to discuss all the relevant information they gathered from the texts they read for the homework reading assignments.

Phase 3: Write the elements for solving the task in keywords – In the third phase, the writer of the discussion group is asked to write down the keywords for solving the tasks. Therefore, the writer is asked to minimize the number of words used to describe the relevant information for their tasks.

Phase 4: (Re-)Write your solution to the task – In the fourth phase, the writer of the discussion group is asked to write the answers to the questions of the task. Only the writer is able to write down the answers to the questions in the TELE, but all other discussion group members help formulating the answers in their assigned roles.

Phase 5: Evaluate your approach as a group – In the fifth phase, the group is asked to evaluate not the task itself, but the approach as a whole group. In this phase, discussion group members have the possibility to write down when the cooperation did not go as planned or what they learned most from each other. An example of the screen of the writer in phase 5 is presented in figure 5.

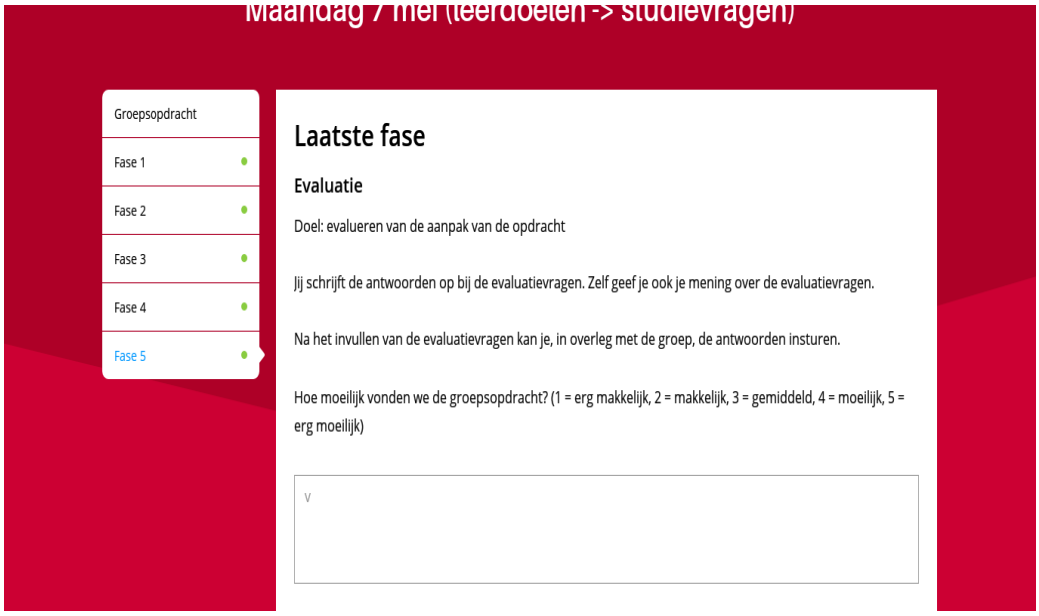

Figure 5: Last of the five phases in which the experimental groups evaluate the tasks.

2.5. Procedure

The procedure of the experiment is visually represented in the schematic figure in Appendix 2 for every course (five courses in total for this research). Each course had a total of one class that contained between 12 and 30 students. Each student was asked to complete a reading test (pre-test) within the application in an examination environment (no talking, no phones, etc.), and was using the same computer during the complete duration of the course. The pre-training reading test consists of texts that were not likely to have ever been read by the students, they were from a major different than the major of the students (all texts for all courses were the same and were from an economics course). After the pre-tests, the students were asked to form four to six discussion groups (the size of the discussion groups was determined by the researcher) of equal size between three and five students in each group (depending on the class size). The discussion groups were randomly assigned to be either in the control group or in the experimental group.

The class was given a homework reading assignment prior to every discussion using the Technology Enhanced Learning Environment. The reading assignments were part of the regular course and contained texts about the major of the student. Every week the student had to go through the same cycle; homework, task and group discussion about the reading assignment. There were six different reading assignments, each was followed by a group discussion, one assignment was given each week:

1. Answering study questions;
2. Making a mindmap;
3. Comparing two texts;
4. Making a summary of a texts;
5. Creating exam questions;
6. Application assignment.

The experimental group was given the same information as the control group, but the information delivery was in different phases during the discussions instead of being presented directly as is done for the control group.

After all the reading assignments, tasks and discussions were done and the course was finished, a new reading test (post-test) was performed, with different texts and questions than the first pre-test, but everything else remained the same. The total information given to both the experimental group and the control group, including the differences between the discussion can be found in appendix 3. The post-training reading test was the same for all courses and were from an economics course, as was the case with the pre-training reading test. The information extracted from both pre- and post-training reading tests were collected, and the raw data was manipulated by the researcher to generate data about the reading behavior of the control and the experimental group.

2.6. Data-analysis

Using SPSS 26.0 for OSX a t-test was performed on the following data as dependent variables; average time reading relevant texts on the reading test, average number of non-relevant texts opened during the reading test, reading test score, the average number of texts read and the average time used to read texts. The group type (control or experimental group) was the independent variable in the t-test (Field, 2015). It was expected to find a significant difference between the experimental and control groups in the post-reading test due the training using phased discussion in the experimental group and not in the pre-training reading test (before the training).

Next to the preparation and facilitation of the experiments and the data analysis, the main contribution of the author to this research was writing a code to manipulate the raw output of the application to generate multiple variables to be used in further research about the reading behavior of students. The manipulation was performed using VBA (Visual Basic for Applications), the author had created a VBA code to extract different variables from the raw data files. The author then preformed a t-test on the data to answer the research questions stated in section 1.3.

3. Results

3.1. Respondents

The research was performed between 13th of September 2018 and 8th of November 2019 and consisted of 104 students. The students were students studying at the University of Applied sciences of Rotterdam that were studying to become a teacher for the secondary education in the Netherlands. The student's data were anonymous therefore, it is not possible for the author to give any information about the specialization of the students, the age or the sex. The students were distributed over different courses, 20 students were studying pharmacology, twelve were studying physics and ICT, 27 were studying educational sociology, 28 were studying research practice and seventeen were studying governmental finances.

One student during the research practice course was switched from experimental to control group right after the pre-training reading test because of incidents within the group. Since this was done before the training started and the reading tests are the same for both the control- and the experimental group this is not expected to influence the outcome of this research. Seventeen students did not make a post-test, they were included in the pre-test and were almost equally dispersed between control and experimental group (65 % were from the experimental group) and were mainly following the educational sociology course (thirteen educational sociology, one pharmacology, one physics and ICT, one research practice and one governmental finances). One student during the research practice course did not make the pre-training reading test but did made the post-training reading test due to absence, the data gathered from this student is only used during the data analyses in the post-training reading test. All available tests were used in the analysis during this master thesis, therefore there were not the same of participants for the pre-training reading test (94 students) as for the post-training reading test (85 students).

3.2. Results per research question

Results will be presented per research question. The gathered data will also be used to answer different research questions by Okkinga and Van Gelderen during further research within the research group of Okkinga and Van Gelderen. Please note that the pre- and post-training reading tests were different from each other: therefore, they cannot be directly compared.

Evaluation of the effect of phased guidance in scripted discussions on task-oriented reading within student groups in higher vocational education

3.2.1. RQ1: Are students who have had phased scripted discussions during training, opening less texts in general before answering a question in the post training reading test than students without phased scripted discussions during training?

To determine if students in both the experimental as the control group had a comparable reading behavior prior to the training in which the experimental group had phased guidance during group discussions and the control group did not have phased guidance during group discussions, a t-test was performed on the average number of texts that were read ($\overline{n_t}$) during the pre-training reading test. During the pre-training reading test the students who would be using phased discussions (experimental group) opened less texts (M = 16.5, SD = 4.2) than the students in the control group who would not be receiving phased discussions (M = 18.1, SD = 5.6), this difference was not significant, $t(94.8) = 1.577$, p = .118. On the post-training reading test after training using the TELE and where the experimental group used the phased discussions, the average number of texts that where read was almost the same for the students in the experimental group (M = 11.6, SD = 4.0) as for students in the control group (M = 11.3, SD = 4.7), the difference was not significant, $t(85) = -.293$, p = .770.

3.2.2. RQ2: Are students who have had phased scripted discussions during training, using more time to read texts before answering a question in the post training reading test than students without phased scripted discussions during training?

To determine if students in both the experimental as the control group had a comparable reading behavior prior to the training in which the experimental group had phased guidance during group discussions and the control group did not have phased guidance during group discussions, a t-test was performed on the average time used to read texts ($\overline{t_t}$) during the pre-training reading test. During the pre-training reading test the students who would be using phased discussions (experimental group) used less time to read texts (M = 149,421 ms, SD = 45,231 ms) than the students in the control group who would not be receiving phased discussions (M = 152,983 ms, SD = 41,739 ms), this difference was not significant, $t(101) = .415$, p = .679. On the post-training reading test after training using the TELE and where the experimental group used the phased discussions, the students in the experimental group used more time to read texts on average (M = 135,227 ms, SD = 54,056 ms) than the students in the control group (M = 124,459 ms, SD = 59,359 ms), the difference was not significant, $t(85) = -.881$, p = .381.

3.2.3. RQ3: Are students who have had phased scripted discussions during training, using more time to read relevant texts before answering a question in the post training reading test than students without phased scripted discussions during training?

To determine if students in both the experimental as the control group had a comparable reading behavior prior to the training in which the experimental group had phased guidance during group discussions and the control group did not have phased guidance during group discussions, a t-test was performed on the average time students were reading relevant texts on the pre-training reading test ($\overline{t_r}$). During the pre-training reading test the students who would be using phased discussions (experimental group) used less time (*M* = 56,637 ms, *SD* = 20,976 ms) reading relevant texts than students who were in the control group (*M* = 62,971 ms, *SD* = 21,678 ms), this difference was not significant, *t*(101) = 1.507, *p* = .135. On the post-training reading test after training using the TELE and where the experimental group used the phased discussions, the average reading time students used to read relevant texts was almost the same for the students in the experimental group (*M* = 48,577 ms, *SD* = 19,660 ms) as for the students in the control group (*M* = 47,197 ms, *SD* = 17,806 ms), the difference was not significant, *t*(85) = -.344, *p* = .732.

3.2.4. RQ4: Are students who have had phased scripted discussions during training, reading less irrelevant texts before answering a question in the post training reading test than students without phased scripted discussions during training?

To determine if students in both the experimental as the control group had a comparable reading behavior prior to the training in which the experimental group had phased guidance during group discussions and the control group did not have phased guidance during group discussions, a t-test was performed on the average number of irrelevant texts read by students on the pre-training reading test ($\overline{n_{nr}}$). During the pre-training reading test the students who would be using phased discussions (experimental group) read less irrelevant texts (*M* = 11.9, *SD* = 3.5) than students who were in the control group (*M* = 12.6, *SD* = 4.5), this difference was not significant, *t*(96.0) = .927, *p* = .356. On the post-training reading test after training using the TELE and where the experimental group used the phased discussions, the average number of irrelevant texts read by students was almost the same for the students in the experimental group (*M* = 7.99, *SD* = 3.18) as for the students in the control group (*M* = 7.53, *SD* = 3.41), the difference was not significant, *t*(85) = -.657, *p* = .519.

3.2.5. RQ5: Are students who have had phased scripted discussions during training, answering the questions on the post training reading test better than students without phased scripted discussions during training?

To determine if students in both the experimental as the control group had a comparable reading behavior prior to the training in which the experimental group would had phased guidance during group discussions and the control group did not have phased guidance during group discussions, a t-test was performed on the fraction of questions that were correctly answered (r_m) during the pre-training reading test. During the pre-training reading test the students who would be using phased discussions (experimental group) answered a smaller fraction of questions correctly (M = .438, SD = .166) than the students in the control group who would be receiving training without phased discussions (M = .455, SD = .155), this difference was not significant, $t(94) = .513$, $p = .609$. On the post-training reading test after training using the TELE and where the experimental group used the phased discussions, the fraction of correctly answered questions was almost the same for the students in the experimental group (M = .609, SD = .179) as for the students in the control group (M = .605, SD = .165), the difference was not significant, $t(85) = -.118$, $p = .906$.

4. Conclusion and discussion

The main research question was: “Do the phased scripted discussions within student groups improve the individual reading behavior of students in higher vocational education?”. The main hypothesis was that students who have had training using the technology-enhanced learning environment with phased scripted discussions will open less irrelevant texts before answering a question, will study relevant text longer and answer questions correctly more frequently than students without phased scripted discussions.

4.1. Conclusions and discussion based on the data

To answer the main research question, four research questions were formulated together with hypothesizes. The first research question (RQ1): “Are students who have had phased scripted discussions during training, opening less texts in general before answering a question in the post training reading test than students without phased scripted discussions during training?” and the hypothesis (H1): “Students who have had phased scripted discussions during the training will open less texts before answering a question on the post training reading test than students without phased scripted discussions during training.” could not be confirmed on the basis of this experiment. There was no significant difference found between the students in the experimental and the students in the control group for the average number of texts read prior to opening a question ($\overline{n_t}$) in both the pre- and post-training reading test.

The second research question (RQ2): “Are students who have had phased scripted discussions during training, using more time to read texts before answering a question in the post training reading test than students without phased scripted discussions during training?” and the hypothesis (H2): “Students who have had phased scripted discussions during the training will use less time reading texts before answering a question on the post training reading test than students without phased scripted discussions during training.” could not be confirmed on the basis of this experiment. There was no significant difference found between the students in the experimental and the students in the control group for the average time span used to read texts prior to opening a question ($\overline{t_t}$) in both the pre- and post-training reading test.

The third research question (RQ3): “Are students who have had phased scripted discussions during training, using more time to read relevant texts before answering a question in the post training reading test than students without phased scripted discussions during training?” and the hypothesis (H3): “Students who have had phased scripted discussions during the training will spent more of their reading time reading relevant texts during the post training reading test than students without phased scripted discussions during training.” could not be confirmed on the basis of this experiment. There was no significant difference found between students in the experimental group and students in the

control group for the average time a relevant text was opened during answering a question ($\overline{t_r}$) in either the pre- or post-training reading test.

The fourth research question (RQ4): "Are students who have had phased scripted discussions during training, reading less irrelevant texts before answering a question in the post training reading test than students without phased scripted discussions during training?" and the hypothesis (H4): "Students who have had phased scripted discussions during the training will read less irrelevant texts before answering a question on the post training reading test than students without phased scripted discussions during training." could not be confirmed on the basis of this experiment. There was no significant difference found between the students in the experimental group and the students in the control group for the average number of occurrences that a non-relevant text was opened during the answering of a question ($\overline{n_{nr}}$) in either the pre- or post-training reading test.

The fifth research question (RQ5): "Are students who have had phased scripted discussions during training, reading less irrelevant texts before answering a question in the post training reading test than students without phased scripted discussions during training?" and the hypothesis (H5): "Students who have had phased scripted discussions during the training will read less irrelevant texts before answering a question on the post training reading test than students without phased scripted discussions during training." could not be confirmed on the basis of this experiment. There was no significant difference found between students in the experimental group when compared with the students in the control group for the fraction of questions that were answered correctly (r_m) in either the pre- or post-training reading test.

Since none of the hypotheses for the research questions could be confirmed, the main hypothesis could not be confirmed. No significant effects could be found between phased discussions and the reading behavior of students. Since they both had discussions and had tasks in which they had to read a lot, it is likely that both groups (independent of whether or not there was phased guidance during the discussions) increased in their ability to read texts and find useful texts faster and answer the questions better on the post test, but this cannot be used as an outcome of this thesis. This experiment is not designed to measure the amount of effect because the two tests were different (in both aspects: texts and questions) and their data can only be used to compare the differences between the two groups (experimental and control group) on either the pre- or post-training reading test. The fact that discussions are helpful to improve reading skills is already confirmed by previous research (Goldenberg, 1992; Perfetti & Roth, 1980).

4.2. Limitations of the research

There were some limitations in the research that could be (partly) the reason for not finding significant differences in the reading behavior of students. Differences between the training for the control and

experimental group were small, this could be a reason for not finding a significant effect on the reading behavior. Since both groups were given the exact same lessons and tasks and were sitting in the same classroom, the situations were the same for both groups, but they could also have had time to communicate with one another (verbal or non-verbal) even though it was not encouraged and students were explicitly told it was undesirable. The difference in the training was based on whether or not they received phased guidance during the scripted discussions and with the other groups sitting in close proximity, students could be tempted to follow each other's discussion with the effect that the groups could (nearly) have had identical discussions. With such a small difference during the discussions, the timeframe of the training could be too short. It is likely that a longer time of phased guidance in scripted guidance during the discussions would increase an effect if there is any effect due to phased guidance in scripted discussions.

The pre- and post-training reading tests were different and could not be directly compared, this is done to prevent students from answering questions on the post-training reading test out of memory from the pre-training reading test. On hindsight a better research method would have been to use and measure the same test twice for both the pre- and post-training reading test, then the results could be used to measure the difference in growth between the control and experimental group. When the two reading tests were equal it could also be possible to measure effects for each question and to monitor changes based on the question types. Because of the design, it was not possible to compare the two tests or questions directly. Since I joined an ongoing research project, I could not change the chosen method.

Another possible explanation for not finding any significant effect is that the way of measuring might not be precise enough. The software collected data based on some assumptions, one is that a student is reading when a text is opened. It is reasonable to think that students sometime lose focus, this could be during the time a text is opened and therefore adding reading time to either a relevant or non-relevant text. Since there were more non-relevant texts it is reasonable that this will affect the reading time of non-relevant texts more than that of relevant texts. It is likely that this effects students with lower reading skills more than students with higher reading skills since students with lower reading skills will likely be more tired (Runyan, 1991).

The reading environment during the tests is not a normal reading environment since students will likely be more used to reading on paper, which can alter the reading behavior of students (Nagy, Anderson, & Herman, 1987). In this study the effect of reading from a digital environment is prevented by giving both reading tests and all assignments on a single device in the same TELE. It is not unlikely that students learn to use this device during the course and will read differently because they are trained in using the device. There is no reason to think this will affect the experimental group dif-

ferently than the control group. The research was conducted using a technology-enhanced learning environment and was therefore, a digital way of presenting texts (Mangen, Walgermo, & Brønnick, 2013). It is likely that students read differently when texts are presented in a digital way rather than an analog way (Dreyer & Nel, 2003). It is possible that this effect of reading different on digital devices is dependent on the reading level of students and therefore, dependent on the effect of the intervention.

With eye tracking, analog texts could be used to monitor and analyze the reading behavior of students (Biedert, Buscher, & Dengel, 2010).To measure the reading behavior of the student more closely to their normal reading environment, in the future a form of eye tracking could be used to find the amount of text that was read without interfering with the reading behavior of the students. Eye tracking is already used in previous research and can be used to closely monitor the focal point of the students (Buscher, Biedert, Heinesch, & Dengel, 2010; Franklin, Smallwood, & Schooler, 2011). Using eye tracking could replace the application and measure the reading behavior more closely while interfering with their reading environment less.

There were only 104 students participating in the research, some effects could be significant if more students were participating (Kanyongo, Brook, Kyei-Blankson, & Gocmen, 2007; Keselman et al., 1998). Besides that, the students were all from different majors which could influence the outcome, since students from different majors likely have differences in their starting reading behavior (Brantmeier, 2002; Raudenbush & Sadoff, 2008). There is also a chance that students from different mayors have different experiences with scripted guided discussions in the past. It is unlikely that the effect of different reading skills or experience with scripted guided discussions between the mayors affected the research group differently than the control group, but it could have increased the variation within both of the groups and therefore, influence the significance of the analysis (Field, 2015). The students that were participating were all in the higher vocational education system, there is a chance that the methods described in this research will affect students in the higher vocational education system differently from students in other education systems.

4.3. Implications on future research and society

This master research is part of a bigger ongoing research by Okkinga. This research gives insight in the effect of the reading tasks on the reading behavior of students. The application to measure the reading level could be used in a follow-up research, the method can extract a lot of data from the reading behavior of students. In a future research the same application could be used to measure the reading behavior of students on an item or question level, to find possible effects of the intervention on the reading behavior during the answering of specific types of questions. It is possible that questions based on differences in knowledge levels will be affected differently by the intervention.

The application to assign roles to students and to integrate all assignments in a single format (with or without the phased discussions) and learning environment is proven useful for the teachers and students. Even when the reading test is not used, the software could be used to help teachers to encourage students to discuss texts and make assignments about the texts that were read. The written script for interpreting the reading behavior during the pre- and the post-training reading test is useful for measuring the reading behavior of students and can be used in further research. The script has been written to work in different environments because general coding was used to ensure that easy adaption can be realized for further research and it can easily be expanded when adjustments are made in the reading tests.

Based on this research no advise can be given on the use of phased guidance in scripted discussions during the training of reading behavior for students in the higher vocational education system, since there are no significant results that give an indication that the phased discussion helps students to improve their reading behavior. Future research should be conducted to find other factors that influence the reading behavior of students and to determine the effect of phased guidance in scripted discussions.

5. References

Afflerbach, P., Pearson, P. D., & Paris, S. G. (2008). Clarifying differences between reading skills and reading strategies. The Reading Teacher, 61(5), 364-373.

Alavi, S. B., & McCormick, J. (2008). The roles of perceived task interdependence and group members' interdependence in the development of collective efficacy in university student group contexts. *British Journal of Educational Psychology, 78*(3), 375-393.

Aronson, E., & Bridgeman, D. (1979). Jigsaw groups and the desegregated classroom: In pursuit of common goals. *Personality and social psychology bulletin, 5*(4), 438-446.

Bailey, N. R., & Scerbo, M. W. (2007). Automation-induced complacency for monitoring highly reliable systems: The role of task complexity, system experience, and operator trust. *Theoretical Issues in Ergonomics Science, 8*(4), 321-348.

Biedert, R., Buscher, G., & Dengel, A. (2010). The eyebook–using eye tracking to enhance the reading experience. *Informatik-Spektrum, 33*(3), 272-281.

Brantmeier, C. (2002). Second language reading strategy research at the secondary and university levels: Variations, disparities, and generalizability. *The Reading Matrix, 2*(3), 1-14.

Brown, A. L., Palincsar, A. S., & Armbruster, B. B. (1984). Instructing comprehension-fostering activities in interactive learning situations. In H. Mandl, N.L. Stein, & T. Trabasso (Eds.), *Learning and comprehension of text*(pp. 255-286). Hillsdale, NJ: Erlbaum.

Buscher, G., Biedert, R., Heinesch, D., & Dengel, A. (2010). Eye tracking analysis of preferred reading regions on the screen. In *CHI '10 Extended Abstracts on Human Factors in Computing Systems*, (pp. 3307-3312). Atlanta, Georgia, USA: Association for Computing Machinery.

Campbell, D. J. (1988). Task complexity: A review and analysis. *Academy of management review, 13*(1), 40-52.

Chesley, H., Cheng, L., Seban, S., & Landau, J. (2005). Scripted text discussion system (U.S. Patent No. 6,795,093). U.S. Patent and Trademark Office. http://patft.uspto.gov/netacgi/nph-Parser?Sect1=PTO2&Sect2=HITOFF&p=1&u=%2Fnetahtml%2FPTO%2Fsearch-bool.html&r=1&f=G&l=50&co1=AND&d=PTXT&s1=%22Scripted+text+discussion+system%22&OS=

Commons, M. L., Trudeau, E. J., Stein, S. A., Richards, F. A., & Krause, S. R. (1998). Hierarchical complexity of tasks shows the existence of developmental stages. *Developmental review, 18*(3), 237-278.

Darabi, A., Arrastia, M. C., Nelson, D. W., Cornille, T., & Liang, X. (2011). Cognitive presence in asynchronous online learning: A comparison of four discussion strategies. *Journal of computer assisted learning, 27*(3), 216-227.

Darisipudi, A. (2006). *Towards a generalized team task complexity model.* (PhD thesis). Retrieved from https://digitalcommons.lsu.edu/cgi/viewcontent.cgi?article=4230&context=gradschool_dissertations

Davis, E. A., & Miyake, N. (2004). Explorations of scaffolding in complex classroom systems. *The Journal of the Learning sciences, 13*(3), 265-272.

De Corte, E., Verschaffel, L., & Van De Ven, A. (2001). Improving text comprehension strategies in upper primary school children: A design experiment. *British Journal of Educational Psychology, 71*(4), 531-559.

Dillenbourg, P. (2002). Over-scripting CSCL: The risks of blending collaborative learning with instructional design. In P. A. Kirschner (Ed.), *Three worlds of CSCL. Can we support CSCL?* (pp. 61-91). Heerlen: Open Universiteit Nederland.

Dillenbourg, P., & Jermann, P. (2007). Designing integrative scripts. In *Scripting computer-supported collaborative learning* (pp. 275-301). New York City: Springer.

Dillenbourg, P., & Tchounikine, P. (2007). Flexibility in macro-scripts for computer-supported collaborative learning. *Journal of computer assisted learning, 23*(1), 1-13.

Expertgroep Doorlopende leerlijnen Taal en Rekenen (2008). Over de drempels met taal. De niveaus voor de taalvaardigheid. Onderdeel van de eindrapportage van de Expertgroep Doorlopende Leerlijnen Taal en Rekenen. Enschede: SLO.

Dreyer, C., & Nel, C. (2003). Teaching reading strategies and reading comprehension within a technology-enhanced learning environment. *System, 31*(3), 349-365.

Durkin, D. (1978). What classroom observations reveal about reading comprehension instruction. *Reading research quarterly*, 481-533.

Edmonds, M. S., Vaughn, S., Wexler, J., Reutebuch, C., Cable, A., Tackett, K. K., & Schnakenberg, J. W. (2009). A synthesis of reading interventions and effects on reading comprehension outcomes for older struggling readers. *Review of educational research, 79*(1), 262-300.

Elderman, F. (2014). Lezen wordt weer leuk! Een praktisch onderzoek naar leesgedrag en-motivatie bij MBO-leerlingen en een persoonlijke aanpak tot verbetering (Master's thesis). Retrieved from https://dspace.library.uu.nl/bitstream/handle/1874/298629/Masterscriptie%3aOnderzoeksverslag%20F.%20Elderman%203343138.pdf?sequence=2&isAllowed=y

Field, A. (2015). *Discovering statistics using IBM SPSS statistics*. Londen: Sage.

Franklin, M. S., Smallwood, J., & Schooler, J. W. (2011). Catching the mind in flight: Using behavioral indices to detect mindless reading in real time. *Psychonomic bulletin & review, 18*(5), 992-997.

Gibbons, P. (2002). *Scaffolding language, scaffolding learning: Teaching second language learners in the mainstream classroom.* Portsmouth, NH: Heinemann.

Goldenberg, C. (1992). Instructional conversations: Promoting comprehension through discussion. *The Reading Teacher, 46*(4), 316-326.

Ham, D.-H., Park, J., & Jung, W. (2011). A framework-based approach to identifying and organizing the complexity factors of human-system interaction. *IEEE Systems Journal, 5*(2), 213-222.

Harlan, M. A., Bruce, C., & Edwards, S. (2016). Inquiry Based Learning Models, Information Literacy, and Student Engagement: A literature review. *School Libraries Worldwide, 22*(2), 23-39.

Harris, T. L., & Hodges, R. E. (1995). The literacy dictionary: The vocabulary of reading and writing: ERIC.

Harvey, C. M., & Koubek, R. J. (2000). Cognitive, social, and environmental attributes of distributed engineering collaboration: A review and proposed model of collaboration. *Human Factors and Ergonomics in Manufacturing & Service Industries, 10*(4), 369-393.

Hermida, J. (2009). The importance of teaching academic reading skills in first-year university courses. The International Journal of Research and Review, 3, 20-28. *https://ssrn.com/abstract=1419247*.

Hu, G., & Li, X. (2017). Asking and answering questions in English-medium instruction classrooms: What is the cognitive and syntactic complexity level? In *English-medium instruction in Chinese universities* (pp. 184-203). Abingdon: Routledge.

Kanyongo, G. Y., Brook, G. P., Kyei-Blankson, L., & Gocmen, G. (2007). Reliability and statistical power: How measurement fallibility affects power and required sample sizes for several parametric and nonparametric statistics. *Journal of Modern Applied Statistical Methods, 6*(1), 9.

Keselman, H. J., Huberty, C. J., Lix, L. M., Olejnik, S., Cribbie, R. A., Donahue, B., ... & Levin, J. R. (1998). Statistical practices of educational researchers: An analysis of their ANOVA, MANOVA, and ANCOVA analyses. Review of educational research, 68(3), 350-386.

Kim, S., & Lombardino, L. J. (2015). Comparing graphs and text: Effects of complexity and task. *Journal of Eye Movement Research, 8*(3):2, 1-17.

Kintsch, W., Patel, V. L., & Ericsson, K. A. (1999). The role of long-term working memory in text comprehension. *Psychologia, 42*(4), 186-198.

Kirkland, M. R., & Saunders, M. A. P. (1991). Maximizing student performance in summary writing: Managing cognitive load. *Tesol Quarterly, 25*(1), 105-121.

Kobbe, L., Weinberger, A., Dillenbourg, P., Harrer, A., Hämäläinen, R., Häkkinen, P., & Fischer, F. (2007). Specifying computer-supported collaboration scripts. *International Journal of Computer-Supported Collaborative Learning, 2*(2-3), 211-224.

Kollar, I., Fischer, F., & Hesse, F. W. (2006). Collaboration scripts–a conceptual analysis. *Educational Psychology Review, 18*(2), 159-185.

Liu, P., & Li, Z. (2012). Task complexity: A review and conceptualization framework. *International Journal of Industrial Ergonomics, 42*(6), 553-568.

Mangen, A., Walgermo, B. R., & Brønnick, K. (2013). Reading linear texts on paper versus computer screen: Effects on reading comprehension. *International journal of educational research, 58*, 61-68.

Mokhtari, K., & Sheorey, R. (2002). Measuring ESL students' awareness of reading strategies. *Journal of developmental education, 25*(3), 2-11.

Nagy, W. E., Anderson, R. C., & Herman, P. A. (1987). Learning word meanings from context during normal reading. *American educational research journal, 24*(2), 237-270.

Okkinga, M. (2018). *Teaching reading strategies in classrooms: does it work?*. University of Twente. https://doi.org/10.3990/1.9789036544481

Okkinga, M., van Steensel, R., van Gelderen, A. J., & Sleegers, P. J. (2018). Effects of reciprocal teaching on reading comprehension of low-achieving adolescents. The importance of specific teacher skills. *Journal of research in reading, 41*(1), 20-41.

Oliveira, I., Tinoca, L., & Pereira, A. (2011). Online group work patterns: How to promote a successful collaboration. *Computers & Education, 57*(1), 1348-1357.

Palinscar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and instruction, 1*(2), 117-175.

Pea, R. D. (2004). The social and technological dimensions of scaffolding and related theoretical concepts for learning, education, and human activity. *The Journal of the Learning sciences, 13*(3), 423-451.

Pearson, P. D., & Duke, N. K. (2002). Effective practices for developing reading comprehension. In A. E. Farstrups & S. J. Samuaels (Eds.), *What reseacrch has to say about reading comprehension*, 3rd ed. (pp. 205-242). Newark, DE: International Reading Association.

Perfetti, C. A., Landi, N., & Oakhill, J. (2005). The acquisition of reading comprehension skill. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook*, (pp. 227-247). Oxford: Blackwell.

Perfetti, C. A., & Roth, S. (1980). Some of the interactive processes in reading and their role in reading skill. In A. M. lesgold & C..A. Perfetti (Eds.), *Interactive processes in reading.* Hillsdale, N.J.: LEA

Raudenbush, S. W., & Sadoff, S. (2008). Statistical inference when classroom quality is measured with error. *Journal of Research on Educational Effectiveness, 1*(2), 138-154.

Rosenshine, B., & Meister, C. (1994). Reciprocal teaching: A review of the research. *Review of educational research, 64*(4), 479-530.

Runyan, M. K. (1991). The effect of extra time on reading comprehension scores for university students with and without learning disabilities. *Journal of learning disabilities, 24*(2), 104-108.

Schellens, T., Van Keer, H., De Wever, B., & Valcke, M. (2007). Scripting by assigning roles: Does it improve knowledge construction in asynchronous discussion groups? *International Journal of Computer-Supported Collaborative Learning, 2*(2-3), 225-246.

Soller, A., Goodman, B., Linton, F., & Gaimari, R. (1998). Promoting effective peer interaction in an intelligent collaborative learning system. *Proceedings of the Fourth International Conference on Intelligent Tutoring Systems (ITS 98)*, San Antonio, TX, 186-195.

Spörer, N., Brunstein, J. C., & Kieschke, U. (2009). Improving students' reading comprehension skills: Effects of strategy instruction and reciprocal teaching. *Learning and instruction, 19*(3), 272-286.

Stegmann, K., Weinberger, A., & Fischer, F. (2007). Facilitating argumentative knowledge construction with computer-supported collaboration scripts. *International Journal of Computer-Supported Collaborative Learning, 2*(4), 421-447.

Strijbos, J.-W., & Fischer, F. (2007). Methodological challenges for collaborative learning research. *Learning and instruction, 17*(4), 389-393.

Strijbos, J.-W., & Weinberger, A. (2010). Emerging and scripted roles in computer-supported collaborative learning. *Computers in Human Behavior, 26*(4), 491-494.

Vaughn, S., Swanson, E. A., Roberts, G., Wanzek, J., Stillman-Spisak, S. J., Solis, M., & Simmons, D. (2013). Improving reading comprehension and social studies knowledge in middle school. *Reading research quarterly, 48*(1), 77-93.

Vervoort, M., & Elffers, L. (2018). Het achterblijvende studiesucces van mbo-instromers in het eerste jaar van het hbo: de rol van vaardigheden, leerstijl, studiekeuze en motivatie. *Tijdschrift voor Hoger Onderwijs, 36*(1), 23-37.

Vidal-Abarca, E., Martinez, T., Salmerón, L., Cerdán, R., Gilabert, R., Gil, L., . . . Ferris, R. (2011). Recording online processes in task-oriented reading with Read&Answer. *Behavior research methods, 43*(1), 179-192.

Vogel, F., Wecker, C., Kollar, I., & Fischer, F. (2017). Socio-cognitive scaffolding with computer-supported collaboration scripts: A meta-analysis. *Educational Psychology Review, 29*(3), 477-511.

Westera, W. J. (2002). *Reciprocal teaching as a school-wide inclusive strategy.* (PhD thesis). Retrieved from http://hdl.handle.net/2292/767

Wise, A. F., Saghafian, M., & Padmanabhan, P. (2012). Towards more precise design guidance: Specifying and testing the functions of assigned student roles in online discussions. *Educational Technology Research and Development, 60*(1), 55-82.

Wolf, C. (2003). iWeaver: Towards' learning style'-based e-learning in computer science education. In 5th *Australasian Computing Education Conference (ACE 2003)*, Adelaide, South Australia, Australia.

Zhang, Y., Li, Z., Wu, B., & Wu, S. (2009). A spaceflight operation complexity measure and its experimental validation. *International Journal of Industrial Ergonomics, 39*(5), 756-765.

Appendix 1

Appendix 1: Relevant text and question comparison

Pre-test relevant text

Table 1: For all the questions (Q1 to Q11 on the top row) on the pre-training reading test there were relevant texts parts, indicated by an x.

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11
Text 1 text part 3	x										
Text 1 text part 14								x			
Text 1 text part 4	x										
Text 1 text part 13				x				x			
Text 2 text part 1											
Text 2 text part 11							x				
Text 2 text part 2										x	
Text 2 text part 4			x								
Text 2 text part 5			x								
Text 2 text part 7					x						
Text 2 text part 8					x						
Text 2 text part 9					x						
Text 3 text part 13	x										
Text 3 text part 3							x				
Text 3 text part 4	x										x
Text 3 text part 6		x									
Text 4 text part 10		x									
Text 4 text part 11		x									
Text 4 text part 12		x									
Text 4 text part 3						x					
Text 4 text part 5		x									
Text 4 text part 6		x									
Text 4 text part 8		x							x		

Post-test relevant text

Table 2: For all the questions (Q1 to Q11 on the top row) on the post training reading test there were relevant texts parts, indicated by an x.

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11
Text 1 text part 1											
Text 1 text part 3											
Text 1 text part 4							x				
Text 1 text part 6					x		x			x	x
Text 2 text part 1											
Text 2 text part 2	x					x					
Text 2 text part 4	x								x		
Text 2 text part 6						x					
Text 2 text part 8						x					
Text 2 text part 9						x					x
Text 3 text part 1											
Text 3 text part 3											
Text 3 text part 5											
Text 3 text part 7			x								
Text 3 text part 9				x							x
Text 4 text part 1			x								
Text 4 text part 3								x			
Text 4 text part 5											
Text 4 text part 7		x						x			
Text 4 text part 8		x						x			
Text 4 text part 10		x									
Text 4 text part 12			x								
Text 4 text part 14									x		x
Text 4 text part 15					x					x	x

Appendix 2: Schematic representation research

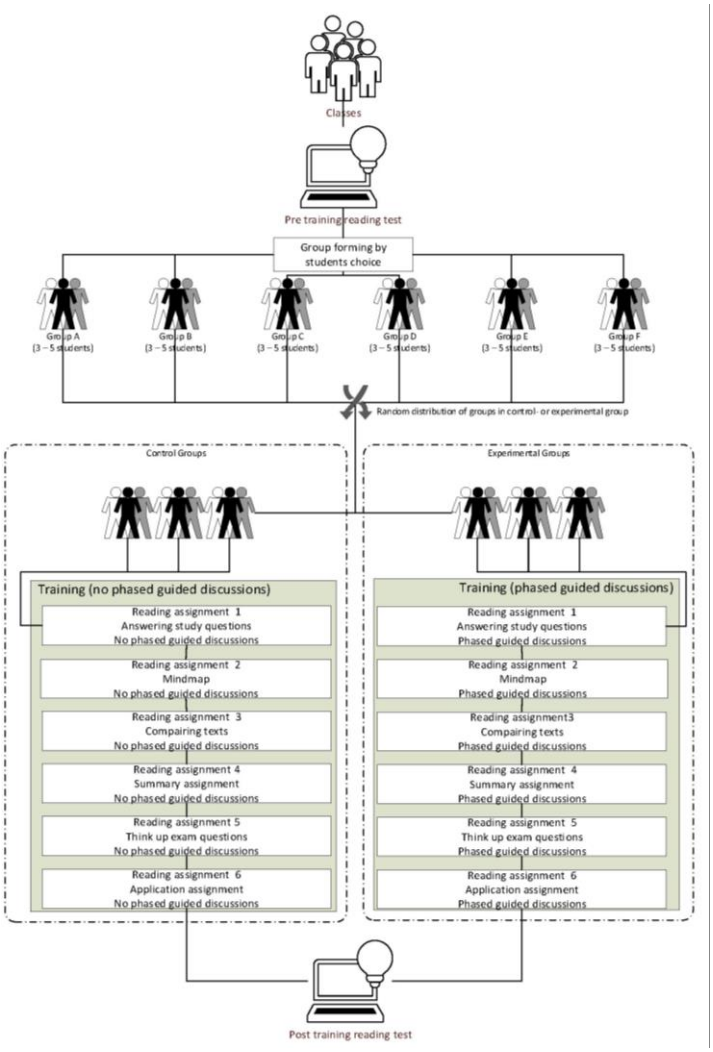

Figure 6:Schematic representation of the research. The schematic shows that each student takes a reading test (pre-training reading test), which is equal for all classes, before being grouped in groups of 3-5 students, these groups were randomly distributed in either in the control- or the experimental group. There were different classes based on the subjects (Pharmacology, Physics and ICT, Educational Sociology, Research Practice and Governmental Finances) which were given at different times and by different teachers. All students regardless of the group (control or experimental group) followed the same reading assignments (but different assignments depending on the subjects for their class). The two groups (experimental and control group) only differ in whether or not they got phased guided discussions during the six different reading assignments. All other steps were the same for both the control and the experimental group. After the reading assignments every individual student took another reading test (post training reading test) which was different from the pre-training reading test but equal for all students.

Appendix 3: Guidance during tasks for experimental group

The text below explains to the students what their roles are and what information is given to the students during their training using the TELE. It starts explaining the different roles to the student and then it gives instructions for the use of the TELE. It starts with the questions of the assignment, as can be seen in figure 7 below. This is just an example since the questions are different for each class and each assignment.

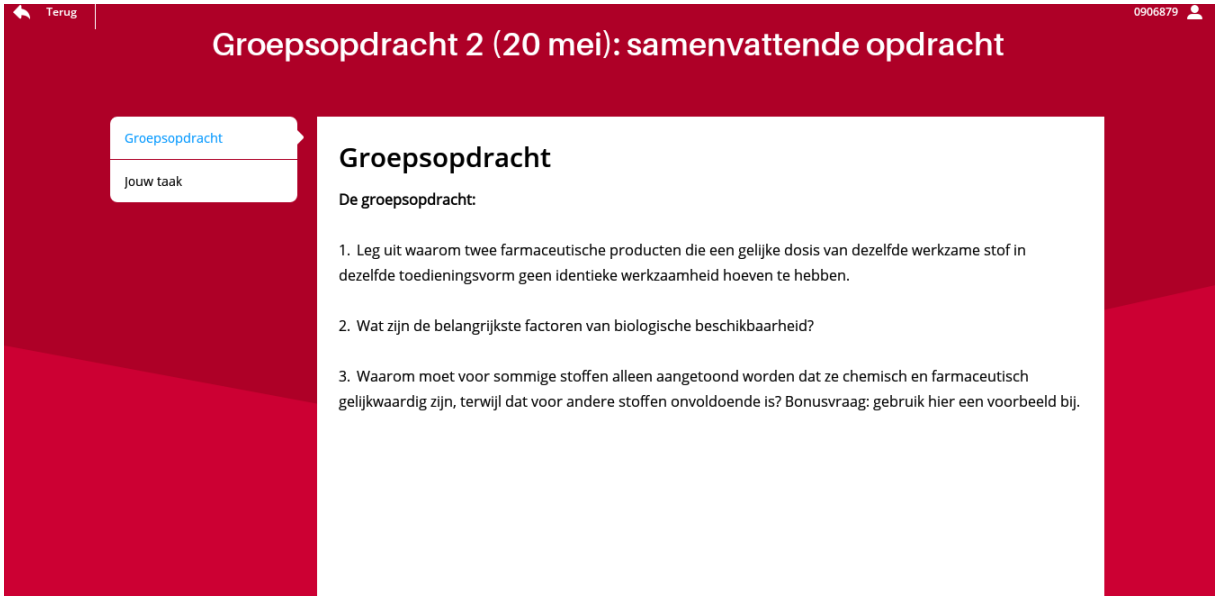

Figure 7: Example of the screen shown to all students with the questions for the assignment (Dutch).

To answer the questions in the assignment, the group should discuss the texts and formulate an answer as a group. For the control group, all students independent of the role receive the same information with the exception that the writer sees a screen to write the answers. The experimental group has different information screens for all roles.

The information below gives more detailed information about the information given to the students through the TELE during the discussions after they have seen the same screen with the assignment (as seen in figure 7).

The first is the information given to all students in the control group, so the chair, the writer and the researcher. The information starts with the explanation of the roles followed by a box to write in for the writer.

Information for the group discussion (control group)

As the chair you are responsible that everything is going smoothly during the group's discussion. As the chair you must monitor the following three tasks:

1. Participation
   Make sure that everybody is participating in the assignment. If somebody is not participating, you are tasked to correct them. This can be done by asking questions like:
   What do you think?, How do you see…?, Do you agree with that?

2. Purposefulness
   Keep the group focused on the goal. Keep track of the assignment en whether a discussion is helpful for the solution to the questions of the assignment. If the group is diverging from the goal of the assignment try to get them back on track.
3. Time monitoring
   Keep track of the time. As a group, you must hand in the assignment. Make sure you finish in time.

The writer takes notes on the discussion and is responsible for writing down the answers to the assignment. If you think the assignment is correctly answered you can send it in. After sending the answers they cannot be changed.

The writer sees the same screen as the chair and thinker but with an empty textbox at the end.

Five phases of the groups discussion for the chair (experimental group)

The experimental group has different screens for all roles and is given in different phases. The phases are described below for the role of chair.

Chair phase 0: Role explanation

As the chair you are responsible that everything is going smoothly during the group's discussion. As the chair you must monitor the following three tasks:

1. Participation
   Make sure that everybody is participating in the assignment. If somebody is not participating, you are tasked to correct them. This can be done by asking questions like:
   What do you think?, How do you see…?, Do you agree with that?
2. Purposefulness
   Keep the group focused on the goal. Keep track of the assignment en whether or not a discussion is helpful for the solution to the questions of the assignment. If the group is diverging from the goal of the assignment try to get them back on track.
3. Time monitoring
4. Keep track of the time. As a group, you must hand in the assignment. Make sure you finish in time.

Chair phase 1: Orientating on the groups assignment

This are the three questions that you should ask the group:

1. What did you find hard about the homework assignment?
2. Why did you find that hard?
3. How hard did you find the groups assignment? (on a scale from 1 to 5)

Make sure everybody answers the questions (including you and the writer). First let the thinkers answer the questions.

Chair phase 2: Exchanging relevant texts for the answering of the assignment

In this phase it is important to find relevant texts that can be used for the assignment. As the chair of the group, you are guiding the process. To keep the discussions going, you can ask questions like:

- Could you give an example of …?
- Why do you think that?
- What is the difference with …?

Ask the thinkers what they think that the most important pieces of texts are for the assignment (they can use their notes).

Find out whether the relevant texts are the same. If they are not the same, what are the differences?

Both you and the writer participate in the exchange of texts.

The writer notes in points the pieces of texts that will be discussed.

Chair phase 3: Selecting relevant points for answering questions.

It is your task that everyone does participate in the discussion about what relevant information will be used to answer the questions.

1. Let the writer tell what points have been noted during the previous phase.
2. Let the team members discuss for every point named by the writer whether that should be used during the answering of the questions.
3. Select as a group what are relevant points and discuss whether there are other relevant points.
4. The writer notes down the relevant points for the answering of the questions.

Chair phase 4: Formulating the final answer

Ask the writer to tell the group what the relevant point are that will be used for the answering of the question (the result of phase 3).

1. Determine the most logic order for the relevant points for answering the question.
2. Formulate a grammatically correct text that contains all the points in the wright order.
3. Check whether the texts could be understood by someone that was not present at the time of the discussion.
4. Check the grammar for correctness.
5. Check whether everyone agrees with the texts. If so, proceed to the next phase.

Chair phase 5: Evaluate the approach

Discuss the following questions with your group:

1. Did you find the assignment hard?
2. If given the same assignment again, would you change something? If so, what would you do different? If not, what approach appeared to work for your group?
3. Discuss who would like what role to get for the next assignment; chair, writer or thinker?
4. Write down you answers of the above questions.

Five phases of the groups discussion for the writer (experimental group)

The writer sees another screen than the chair, the information for the writer in the experimental group is given in phases and are explained below. The writer will see textboxes in which it will answer the questions (based on the instructions by the chair of the group). The textboxes are written in italic, and if information is shown out of previous phases it is also written in italic.

Writer phase 0: Role explanation

As the writer you are responsible for finishing, minute and formulating answers.

Writer phase 1: Orientating on the groups assignment

Listen to the instructions of chair of the group.

Writer phase 2: Exchanging relevant texts for the answering of the assignment

Listen to the instructions of the chair of the group.

Note points of the pieces of texts that will be discussed in the following phase.

Empty textbox.

Writer phase 3: Selecting relevant points for answering questions.

Listen to the instructions of the chair of the group.

Note down the relevant points for the answering of the questions.

Textbox that contains text of phase 2.

Writer phase 4: Formulating the final answer

Listen to the instructions of the chair of the group.

Write down the texts that is the final answer.

Text box that contains text of phase 3.

Writer phase 5: Evaluate the approach

Listen to the instructions of the chair of the group.

Did you find the assignment hard?

Empty textbox

If given the same assignment again, would you change something? If so, what would you do different? If not, what approach appeared to work for your group?

Empty textbox

Discuss who would like what role to get for the next assignment; chair, writer or thinker?

Empty textbox

Five phases of the groups discussion for the thinker (experimental group)

The thinker in the experimental group does not get a lot of information, but mainly listens to the chair of the group. The information shown on screen during the different phases of the discussions is shown below.

Thinker phase 0: Role explanation

As the thinker you have an important role to participate in the discussion. It is important that you have pieces of texts that you deem important for the answering of questions during the assignment.

Thinker phase 1: Orientating on the groups assignment

Listen to the instructions of chair of the group.

Thinker phase 2: Exchanging relevant texts for the answering of the assignment

Listen to the instructions of the chair of the group.

Thinker phase 3: Selecting relevant points for answering questions.

Listen to the instructions of the chair of the group.

Thinker phase 4: Formulating the final answer

Listen to the instructions of the chair of the group.

Thinker phase 5: Evaluate the approach

Listen to the instructions of the chair of the group.